AFOSR 70-1966 TR

SPMA

AD709226

RECURSIVE FILTERING FOR SYSTEMS WITH SMALL BUT NONNEGLIGIBLE NONLINEARITIES

H.W. Sorenson,* A.R. Stubberud +

* Currently: Guest Scientist at. Institut für Steuer- und Regeltechnik, Deutsche Versuchsanstalt für Luft- und Raumfahrt, Oberpfaffenhofen, West Germany Formerly: University of California, Los Angeles.

+ Assistant Professor of Engineering, University of California, Los Angeles

This research was sponsored by Grant AFOSR 699-67

Submitted on 27.6.6

20

ABSTRACT

Linear estimation theory has been applied extensively to nonlinear systems by assuming that perturbations from a reference solution can be described by linear equations. As long as the second order (and higher) terms in the perturbation equations are negligible, linear estimation techniques have been found to yield satisfactory response. Many examples have been encountered in which the linear theory is not satisfactory, however, and it is to this situation that attention is directed here. Time-discrete systems in which the second order effects are small but nonnegligible are considered. Recursion relations for the conditional mean and covariance are developed. While these relations yield approximations to the true values of these moments, they are superior to the approximations provided by applying linear theory to a nonlinear system. Some results for a simple system are presented in which the response from linear and nonlinear filters is compared.

I. INTRODUCTION

Considerable attention has been devoted to the problem of determining unbiased, minimum variance estimates of the state of linear systems from noisy measurement data. For linear systems with additive white noise the solution, frequently referred to as the Kalman filter, is well-known [1] and it has received application to a variety of engineering problems. When the system is nonlinear, linear perturbation theory is introduced and the linear filter is applied to estimate the state perturbations. This has been found to provide satisfactory results in most cases but occasions [2] have been encountered in which nonlinear effects prove to have a very deleterious effect upon the filter response. In the succeeding discussion a nonlinear filter is presented that is appropriate for use when the second order terms are small but nonnegligible. Computational results indicate that a significant improvement relative to a linear filter can be obtained using these relations.

A considerable portion of the recent research into nonlinear filtering has been concerned with the determination of the a posteriori density function of the state conditioned on all available data. Having this density, then any type of estimate (e.g. minimum mean square error, minimum absolute, most probable) can be obtained, theoretically. Because of the difficulties involved in using the general results that have been obtained, attention has been directed to methods of approximating either the density or its moments. This approach has been taken by Bass, Norum and Schwartz [3] for continuous systems. They approximate the true a posteriori density by assuming that it is represented by the first and second order moments (all higher order moments are negligible). In the investigations presented below, time-discrete systems are considered and the a posteriori density is assumed to be gaussian and recursion relations for the mean and covariance are derived. The relations that are obtained are similar to the relations for linear systems but incorporate the effect of second order terms in the plant and measurement equations.

The general problem and its solution are presented and discussed in Section 2. No justification for the results are given there. This discussion is given in Section 3 for a scalar system. A scalar rather than vector system is considered because the notation is simpler and thereby avoids uninteresting complications in the discussion. The generalization to the multidimensional results presented in Section II is straight forward. A simple numerical example is discussed in Section IV and the response of the nonlinear and linear filters is compared.

II. GENERAL PROBLEM AND ITS SOLUTION

The state x_k of a dynamical system is to be estimated from a collection of discrete measurement data $(z_1, \ldots z_k)$. The evolution of the state is assumed to be described by a system of nonlinear difference equations

$$x_k = F_k x_{k-1} + g_k(x_{k-1}) + w_{k-1} \tag{2.1}$$

The m-dimensional measurement data z_k at each sampling time have a known nonlinear relationship with the state

$$z_k = H_k x_k + e_k(x_k) + v_k \tag{2.2}$$

In (2.1) and (2.2) the system is assumed to be completely known except for the initial state x_o and the plant and measurement noise sequences w_j and v_j. The n and m-dimensional vector functions g_k and e_k are known and are such that the i^{th} component is given by

$$g_k^i = x_{k-1}^T G_k^i x_{k-1} \quad (i = 1,2, \ldots,n)$$

$$e_k^i = x_k^T E_k^i x_k \quad (i = 1,2, \ldots,m)$$

where the G_k^i and E_k^i are symmetric (n x n) matrices. The F_k is a (n x n) matrix and the H_k is a (m x n) matrix.

The initial state and the noise sequences are assumed explicitly to be gaussian and to be mutually uncorrelated. Further, the noise sequences

are taken to be uncorrelated between sampling times and shall be referred to as white noise sequences. The initial state has the density

$$p(x_o) = k_o \exp - \tfrac{1}{2}(x_o-a)^T M_o^{-1} (x_o-a) \tag{2.3}$$

and at each sampling time the noise samples have the densities

$$p(w_j) = k \ \exp - \tfrac{1}{2} \ w_j^T \ Q_j^{-1} \ w_j \tag{2.4}$$

$$p(v_j) = k_v \ \exp - \tfrac{1}{2} \ v_j^T \ R_j^{-1} \ v_j \tag{2.5}$$

There are many possible types of estimates that could be considered for this problem. Cox [4] has considered estimates for nonlinear time-discrete systems that are chosen to maximize the a posteriori density function. The conditional mean is known to give the estimate that minimizes the mean square error and this estimate is considered here. Thus, the discussion will be directed toward approximating the mean of the a posteriori density function in a recursive fashion.

The manner in which the a posteriori density $p(x_k/z^k)$ evolves* from one sampling time to the next is easily established. Specifically when the noise is uncorrelated between sampling times, then the general recursion relations are given by

$$p(x_k/z^k) = \frac{p(x_k/z^k) \ p(z_k/x_k)}{p(z_k/z^{k-1})} \tag{2.6}$$

and

$$p(x_k/z^{k-1}) = \int p(x_{k-1}/z^{k-1}) \ p(x_k/x_{k-1}) dx_{k-1} \tag{2.7}$$

where the normalizing constant $p(z_k/z^{k-1})$ is

$$p(z_k/z^{k-1}) = \int p(x_k/z^{k-1}) \ p(z_k/x_k) dx_k \tag{2.8}$$

---

*The notation a^k will be used to denote the set $(a_1, a_2, \dots, a_k)$

The integrals are n-dimensional in (2.7) and (2.8).

The density $p(x_k/z^{k-1})$ described by (2.6) will be referred to as the prediction density whereas $p(x_k/z^k)$ will be designated as the filtering density. In general it is very difficult to perform in a closed form the operations required in (2.7) - (2.8). Also, it is difficult to determine the moments (specifically the mean) from (2.6). This is discussed in Reference 5. However, if the moments of $p(x_{k-1}/z^{k-1})$ are known, it is a relatively straight forward procedure to determine the moments of the prediction density $p(x_k/z^{k-1})$ if the plant noise is additive and gaussian. This will be discussed further in Section III as will the difficulties involved in obtaining the moments of $p(x_k/z^k)$.

When the system is linear (i.e. g_k and e_k are identically zero in (2.1) and (2.2)), then the operations in (2.6) - (2.8) can be carried exactly and are known to be

$$p(x_k/z^{k-1}) = k_{k/k-1}\exp - \frac{1}{2}(x_k-\hat{x}_{k/k-1})^T P_{k/k-1}^{-1}(x_k-\hat{x}_{k/k-1}) \qquad (2.9)$$

and

$$p(x_k/z^k) = k_k\exp - \frac{1}{2}(x_k-\hat{x}_k)^T P_k^{-1}(x_k-\hat{x}_k) \qquad (2.10)$$

where

$$\hat{x}_{k/k-1} = F_k\, x_{k-1} \qquad (2.11a)$$

$$P_{k/k-1} = F_k\, P_{k-1}\, F_k^T + Q_{k-1} \qquad (2.11b)$$

and

$$\hat{x}_k = \hat{x}_{k/k-1} + K_k[z_k-H_k\hat{x}_{k/k-1}] \qquad (2.12a)$$

$$K_k = P_{k/k-1}\, H_k^T[H_k P_{k/k-1}\, H_k^T + R_k]^{-1} \qquad (2.12b)$$

$$P_k = P_{k/k-1} - K_k\, H_k\, P_{k/k-1} \qquad (2.12c)$$

Equations (2.11)- (2.12) are known as the Kalman filter equations. These relations have been stated for ease of reference for the following discussion. When the system described by (2.1) - (2.2) is considered, it is no longer possible to obtain the exact relations for the moments. Because of the nonlinearities, the prediction and filtering densities will lose the gaussian character. In general, these densities will lose their symmetric properties and the skewness will be nonzero.. This means that the maximum point of the density will no longer correspond to the conditional mean.

If the nonlinearities are relatively small, then the skewness will be small and the density will retain an essentially gaussian character. Supposing that the $p(x_k/z^{k-1})$ and $p(x_k/z^k)$ can be approximated by the gaussian density even in the presence of the plant and measurement nonlinearities, then these densities will be described by (2.9) - (2.10). The moments $\hat{x}_{k/k-1}$, $P_{k/k-1}$, $\hat{x}_k$, P_k will not correspond with (2.11) - (2-12) in this case because of the g_k and e_k in (2.1) - (2.2). The following relations are proposed as approximations of these moments. Their derivation is discussed in Section III.

$$\hat{x}_{k/k-1} = F_k x_{k-1} + \hat{g}_k \tag{2.13a}$$

$$P_{k/k-1} = \left[F_k+2\Phi_k\right] P_{k-1} \left[F_k+2\Phi_k\right]^T + Q_{k-1} +2\Theta_k \tag{2.13b}$$

where the i^{th} component of $\hat{g}_k$ is

$$(\hat{g}_k)_i = \hat{x}_{k-1}^T G_k^i \hat{x}_{k-1} + \text{trace } G_k^i P_{k-1}$$

The i^{th} row of the matrix Φ_k is

$$(\Phi_k)_i = \hat{x}_{k-1}^T G_k^i$$

and the ij^{th} element of the matrix Θ_k is

$$(\Theta_k)_{ij} = \text{trace } (G_k^i P_{k-1} G_k^j P_{k-1})$$

For the filtering density the moments are approximated by

$$\hat{x}_k = \hat{x}_{k/k-1} + K_k \left[z_k - \hat{z}_{k/k-1}\right] \tag{2.14a}$$

$$P_k = P_{k/k-1} - K_k \left[H_k + 2\Psi_k\right] P_{k/k-1} \tag{2.14b}$$

where

$$K_k = P_{k/k-1} (H_k+2\Psi_k)^T \left[(H_k+2\Psi_k)P_{k/k-1} (H_k+2\Psi_k)^T + R_k + 2\Delta_k\right]^{-1}$$

$$\hat{z}_{k/k-1} = H_k \hat{x}_{k/k-1} + \hat{e}_k$$

The i^{th} component of $\hat{e}_k$ is

$$(\hat{e}_k)_i = \hat{x}_{k/k-1}^T E_k^i \hat{x}_{k/k-1} + \text{trace } E_k^i P_{k-1}$$

and the i^{th} row of Ψ_k is

$$(\Psi_k)_i = \hat{x}_{k/k-1}^T E_k^i$$

The ij^{th} element of the matrix Δ_k is

$$(\Delta_k)_{ij} = \text{trace } (E_k^i P_{k/k-1} E_k^j P_{k/k-1})$$

Note that if there are no nonlinear terms, then (2.13) and (2.14) reduce immediately to (2.11) - (2.12).

It is interesting to examine these relations by comparing (2.13) - (2.14) with (2.11) - (2.12). First, note that the nonlinearities have the effect of modifying the linear transition matrix F_k in (2.13b) and the linear observation matrix H_k in (2.14b). In the former the F_k of the linear filter is replaced by $(F_k+2\Phi_k)$ and in the latter the H_k is modified to $(H_k+2\Psi_k)$. This provides the primary influence of the nonlinearity upon the covariance matrix. Further, there is an additional effect which can be viewed as an increase in the plant and measurement noise covariance matrices. The plant noise covariance Q_{k-1} is increased to $(Q_{k-1}+2\Theta_k)$ and the measurement noise covariance becomes $(R_k+2\Delta_k)$.

III. DERIVATION OF THE APPROXIMATING MOMENTS

In this Section the derivation of the moments stated in Section II is discussed in terms of a scalar system. The scalar system is discussed here in order to simplify the notation. The extension to the general multidimensional results is straight forward.

Consider the prediction density first. The only assumption required to establish $\hat{x}_{k/k-1}$ and $p^2_{k/k-1}$ is that the a posteriori density $p(x_{k-1}/z^{k-1})$ is gaussian at each sampling time. The scalar version of (2.1) can be written as

$$x_k = f_k x_{k-1} + g_k x^2_{k-1} + w_{k-1} \tag{3.1}$$

The density for the noise sequence will be written as

$$p(w_k) = k_w \exp - \frac{1}{2} \left(\frac{w_k}{q_k}\right)^2$$

The moments of $p(x_k/z^{k-1})$ can be obtained from (2.7) by noting that

$$E\left[x_k^i/z^{k-1}\right] = \int x_k^i \; p(x_k/z^{k-1}) dx_k$$

Substituting (2.7) and iterating integrals, one obtains

$$= \int dx_k \; p(x_{k-1}/z^{k-1}) \int x_k^i \; p(x_k/x_{k-1}) dx_k \tag{3.2}$$

The density $p(x_k/x_{k-1})$ is known from (3.1) and (2.4). It is

$$p(x_k/x_{k-1}) = k_w \exp - \frac{1}{2} \left(\frac{x_k - f_k x_{k-1} - g_k x^2_{k-1}}{q_{k-1}}\right)^2 .$$

Using this relation, the innermost integral in (3.2) is easily determined. In the case when $i=1$, the mean value is

$$E\left[x_k/x_{k-1}\right] = f_k \; x_{k-1} + g_k x^2_{k-1}$$

Then, from the gaussian property for $p(x_{k-1}/z^{k-1})$, one obtains

$$E\left[x_k/z^{k-1}\right] = E\left[f_k x_{k-1} + g_k x_{k-1}^2 /z^{k-1}\right]$$

$$= f_k \hat{x}_{k-1} + g_k E\left[x_{k-1}^2/z^{k-1}\right]$$

$$\hat{x}_{k/k-1} = f_k \hat{x}_{k-1} + g_k(p_{k-1}^2 + \hat{x}_{k-1}^2) \tag{3.3}$$

This is the scalar version of (2.13a). It can be obtained without recourse to the general Bayesian relation (2.7) simply by computing the conditional expectation of x_k using (3.1).

The error in this estimate is given by

$$x_k - \hat{x}_{k/k-1} = f_k(x_{k-1}-\hat{x}_{k-1}) + g_k\left[(x_{k-1}-\hat{x}_{k-1})^2 + 2(x_{k-1}-\hat{x}_{k-1})\hat{x}_{k-1}-p_{k-1}^2\right] + w_{k-1}$$

Using this result, the variance is

$$E\left[(x_k-\hat{x}_{k/k-1})^2/z^{k-1}\right] \overset{Df}{=} p_{k/k-1}^2$$

$$= (f_k+2g_k \hat{x}_{k-1})^2 p_{k-1}^2 + 2g_k^2 p_{k-1}^4 + q_{k-1}^2 \tag{3.4}$$

Equation (3.4) is the scalar version of (2.13b). Note that the only approximation of $\hat{x}_{k/k-1}$ and $p_{k/k-1}^2$ is that the $p(x_{k-1}/z^{k-1})$ is gaussian. The nonlinearities destroy the gaussian character of the density however so (3.3) and (3.4) will be approximations.

The effect of the plant nonlinearity upon the density can be studied by determining the third and fourth central moments. If the density $p(x_k/z_{k-1})$ were gaussian, then the third central moment would be identically zero and the fourth central moment would be $3p_{k/k-1}^4$. Consider what the values actually become. The third central moment is found to be

$$E\left[(x_k-\hat{x}_{k/k-1})^3/z^{k-1}\right] = 2g_k\, p_{k-1}^4\left[3(f_k+2g_k\hat{x}_{k-1})^2-g_k^2 p_{k-1}^2\right] \tag{3.5}$$

and the fourth central moment is

$$E\left[(x_k-\hat{x}_{k/k-1})^4/z^{k-1}\right] = 3p_{k/k-1}^4 + 10g_k^4\, p_{k-1}^8 + 4g_k^2(f_k+2g_k\,\hat{x}_{k-1})^2 p_{k-1}^6 \tag{3.6}$$

From (3.5) it is clear that the presence of the nonlinear coefficient g_k will cause the third central moment to be nonzero. Similarly, the g_k causes the fourth central moment to change from the gaussian value. The changes from the gaussian moments are seen to be proportional to p_{k-1}^4 in (3.5) and p_{k-1}^6 in (3.6). If the filter converges to the true value of the state as the number of samples becomes large, one sees then that the gaussian-destroying terms converge to zero much more rapidly than does the variance p_k^2. This implies that the gaussian approximation will improve as k becomes large.

The derivation of the moments for the prediction density is not difficult. This is not the case for the moments of the filtering density. Consider for a moment the calculation of $p(x_k/z^{k-1})$ from (2.6). First, it is clear that knowledge of $p(x_k/z^{k-1})$ and the measurement equation

$$z_k = h_k\, x_k+e_k x_k^2+v_k \tag{3.7}$$

defines $p(x_k/z^k)$ without error. The density for the v_k's is assumed to be $p(v_k) = k_v \exp - \frac{1}{2}(\frac{v_k}{r_k})^2$. That is, the $p(x_k/z^k)$, using (2.6), is formed by an algebraic combination of the densities. Unfortunately, it is not an easy task to determine the moments of the density that results. To understand this, note that

$$p(z_k/x_k) = k_v \exp - \frac{1}{2}\left(\frac{z_k-h_k x_k-e_k x_k^2}{r_k}\right)^2 \tag{3.8}$$

Assuming that $p(x_k/z^{k-1})$ is gaussian, the $p(x_k/z^k)$ can be computed from (2.6) and can be written as

$$p(x_k/z^k) = k_N \exp-\frac{1}{2}\left(\frac{x_k-\hat{\xi}_k}{\pi_k}\right)^2 \exp - \frac{1}{2}\left(\frac{-2e_k z_k x_k^2+2h_k e_k x_k^3+e_k^2 x_k^4}{r_k^2}\right) \tag{3.9}$$

where

$$\frac{1}{\pi_k^2} = \frac{h_k^2}{r_k^2} + \frac{1}{p_{k/k-1}^2}$$

$$\hat{\xi}_k = \pi_k^2 \left(\frac{h_k z_k}{r_k^2} + \frac{\hat{x}_{k/k-1}}{p_{k/k-1}^2} \right)$$

The k_N is a quantity not involving the state x_k. The π_k^2 and $\hat{\xi}_k$ are the variance and mean that is obtained when the system is linear. Thus, the measurement nonlinearity modifies the filtering density for linear systems by appearing as an exponential factor. This factor contains the fourth power of the state. It is apparent that the moments $E[x_k^i/z^k]$ cannot be computed directly from (3.9) because of the nature of the second exponential factor. Thus, it becomes necessary to approximate the moments in some way. This problem has been discussed in Reference 5.

For this discussion the approximation shall be accomplished indirectly. First, one sees that the moments relating to the measurement (i.e. $E\left[z_k^i/z^{k-1}\right]$) can be computed in the same way as was done for the prediction density. Thus, one obtains

$$\hat{z}_{k/k-1} = h_k \hat{x}_{k/k-1} + e_k (p_{k/k-1}^2 + \hat{x}_{k/k-1}^2) \tag{3.10}$$

When a measurement z_k is processed, the error in the estimate $\hat{z}_{k/k-1}$ is

$$\epsilon_k = z_k - \hat{z}_{k/k-1}$$

Then, hypothesize that an estimate $\hat{x}_k$ is obtained by adding the error ϵ_k linearly to the predicted estimate $\hat{x}_{k/k-1}$

$$\hat{x}_k = \hat{x}_{k/k-1} + K_k (z_k - \hat{z}_{k/k-1}) \tag{3.11}$$

The gain matrix K_k shall be determined so that the mean square error

$$E_{z_k}\left[E(x_k-\hat{x}_k)/z^{k-1}\right] \text{ is minimized.}$$

Note the averaging is restricted to the measurement z_k. The error is found to be

$$x_k-\hat{x}_k = (x_k-\hat{x}_{k/k-1}) - K_k\Big[h_k(x_k-\hat{x}_{k/k-1})+e_k\,(x_k-\hat{x}_{k/k-1})^2+2\hat{x}_{k/k-1}(x_k-\hat{x}_{k/k-1}) - p^2_{k/k-1}\Big] + v_k$$

It follows that the K_k that accomplishes the minimization is given by

$$K_k = \frac{(h_k+2g_k\,\hat{x}_{k/k-1})p^2_{k/k-1}}{\left[(h_k+2g_k\,\hat{x}_{k/k-1})^2\,p^2_{k/k-1}+2g_k\,p^4_{k/k-1}+r^2_k\right]} \tag{3.12}$$

The variance is approximated by $E_{z_k}\left\{E\left[(x_k-\hat{x}_k)^2/z^{k-1}\right]\right\}$ and is found to be

$$p^2_k = p^2_{k/k-1} - K_k(h_k+2e_k\hat{x}_{k/k-1})p^2_{k/k-1} \tag{3.13}$$

Equations (3.10) - (3.13) are the scalar versions of (2.13) - (2.14). Although these results are admittedly obtained in a suboptimal fashion, it is worth observing that this procedure yields the optimal solution for linear systems [1]. Note that the estimate is stagewise unbiased. That is, one finds that $E_{z_k}[\hat{x}_k]$ is equal to $E[x_k]$ when $E[x_{k-1}] = \hat{x}_{k-1}$. Computational results obtained from these filter equations are encouraging and suggest that the approximation is reasonably accurate and certainly is an improvement over a purely linear approximation.

IV. A NUMERICAL EXAMPLE

In this section a simple problem is considered in order to compare the linear and nonlinear filters. The example has been discussed previously by Denham and Pines [2] in connection with the inadequacy of a linear filter when the nonlinearity is comparable to the measurement noise.

Consider a scalar system with a static plant

$$x_k = x_{k-1} \tag{4.1}$$

and measurements described by

$$z_k = x_k^2 + v_k \tag{4.2}$$

The statistics for the initial state and measurement noise are

$$E[x_o] = a \; ; \quad E[(x_o - a)^2] = m_o^2$$

$$E[v_k] = 0 \; ; \quad E[v_k^2] = r_k^2$$

The state x_k will be estimated by using perturbation theory. Then, the results of the preceding sections will be applied to estimate the perturbations. The perturbation equations are

$$\delta x_k = \delta x_{k-1} \tag{4.3}$$

$$\delta z_k = 2x_{k-1}^* \delta x_k + 2\,\delta x_k^2 + v_k \tag{4.4}$$

where x_{k-1}^* is the nominal state and

$$\delta x_k \overset{Df}{=} x_k - x_k^*$$

The nominal was chosen to be

$$x_o^* = a$$

and this value was retained for all k.

This system was simulated on a digital computer. The initial perturbation δx_o and the measurement noise sequence were obtained from a Gaussian random number generator according to the prescribed statistics. Some typical results are portrayed in Figures 1 and 2. In these Figures the standard deviation p_k for the nonlinear filter is plotted as is the absolute value of the error $|x_k - \hat{x}_k|$ for the linear and nonlinear filter. The p_k for the nonlinear filter is given because it is somewhat larger than for the linear filter but is not significantly different. In Figure 1 the variance of the initial perturbation is 0.01 and this was increased to 0.1 for the results shown in Figure 2. The comparison of p_k and $|x_k - \hat{x}_k|$ indicates that the linear filter behaves somewhat satisfactorily for the smaller initial deviation (i.e. Figure 1) although it seems to be diverging as the number of samples increases. In the second case the linear filter is obviously inadequate because the error and its statistic are in complete disagreement. On the other hand the nonlinear filter gives significantly better results and suggests that a great deal more confidence can be placed in the estimates. These results are typical of those that have been obtained. (Unfortunately, the data from a complete Monte Carlo simulation cannot be presented at this time).

The numerical results that have been obtained suggest several conclusions. The two most important are stated below.

(1) Unless the measurement noise is "small", the linear filter and the nonlinear filter give essentially the same response. A precise definition of "small" shall not be attempted other than to say that the noise must be small relative to the second order effects.

(2) When differences between the linear and nonlinear filters do arise, the latter gives consistently better results in the sense that the error and the statistic are consistent with one another.

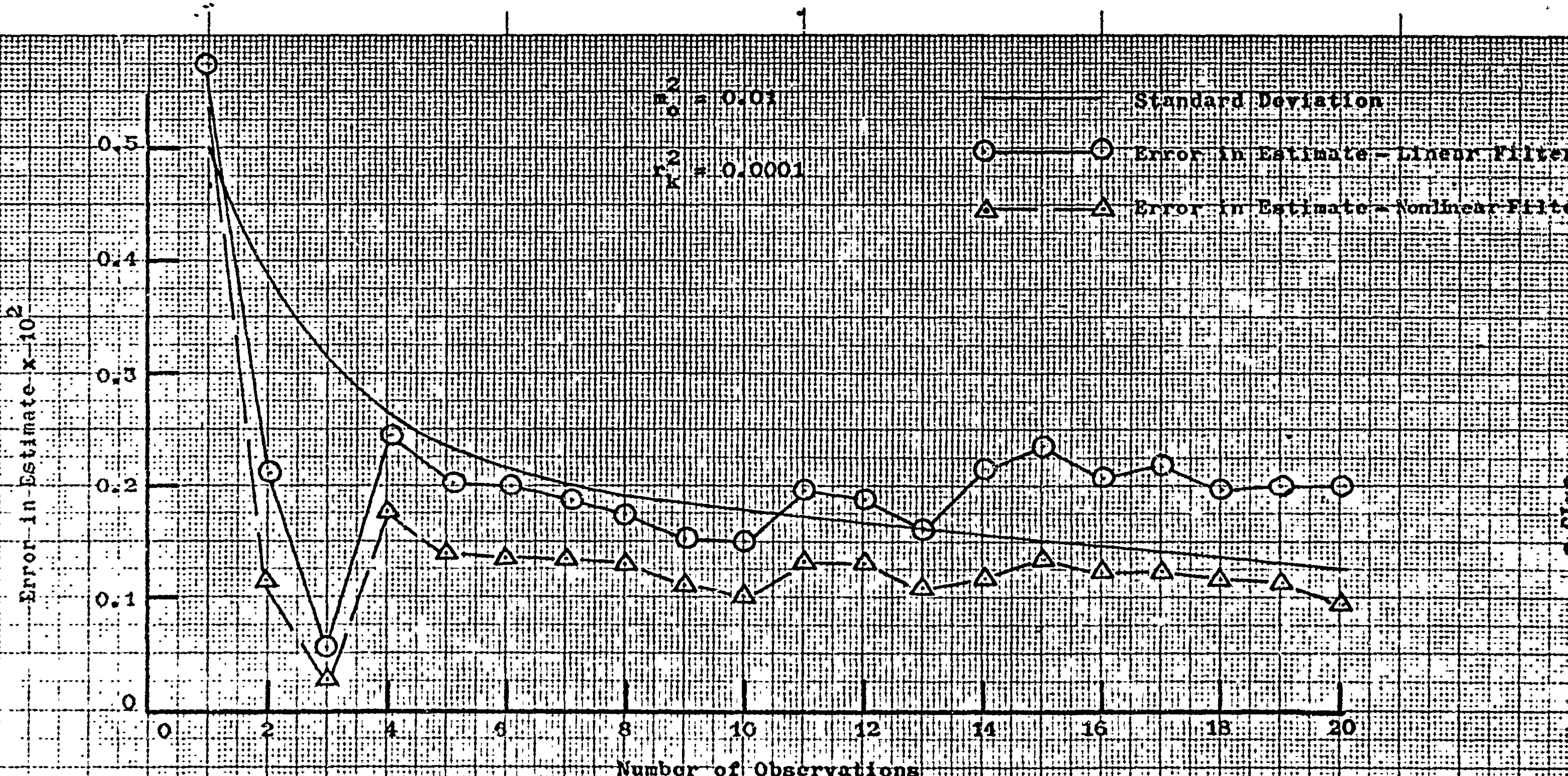

Figure 1; Comparison of Filter Response - Small Initial Perturbation

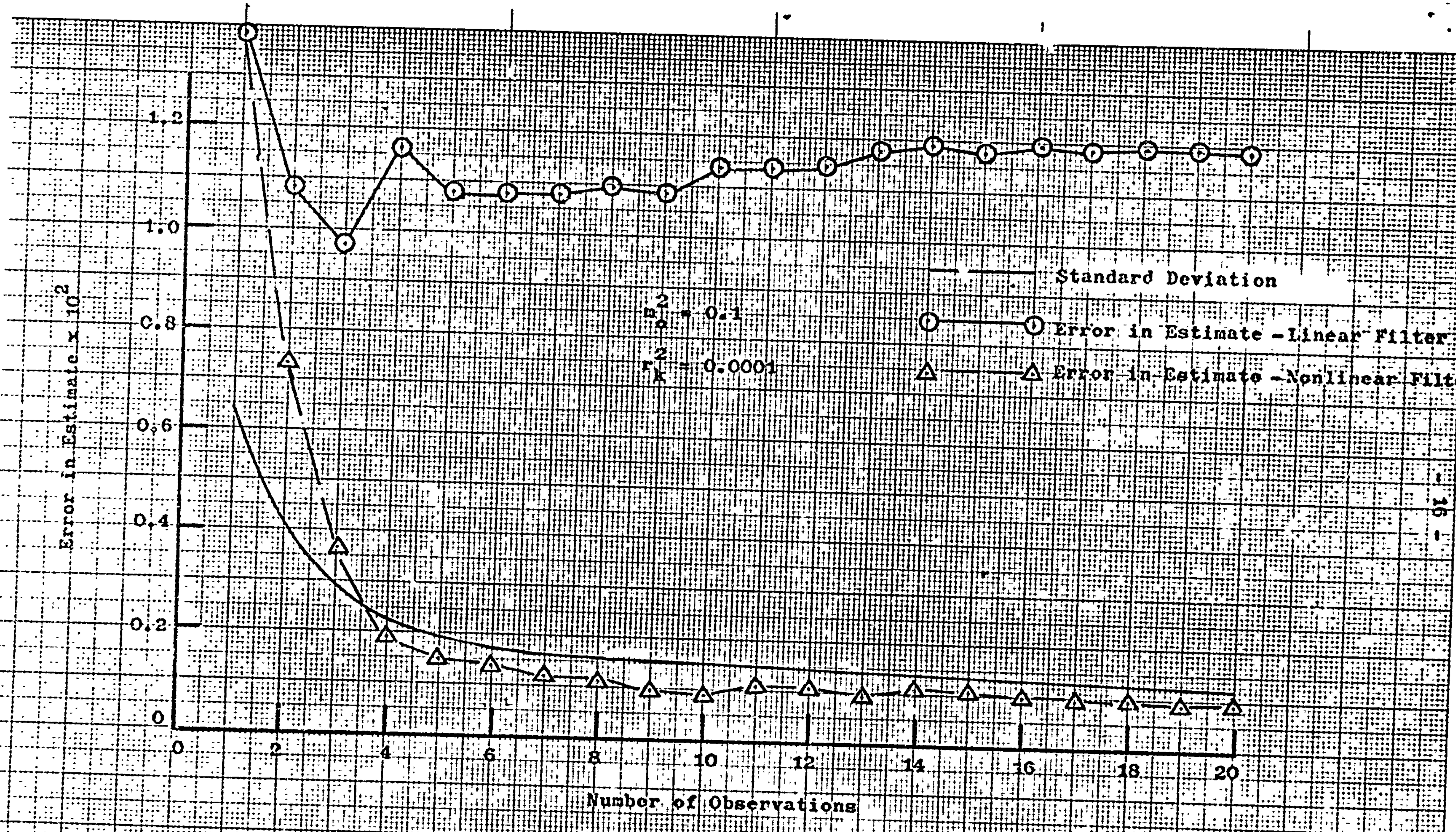

Figure 2: Comparison of Filter Response - Large Initial Perturbation

V. CONCLUSIONS

The problem of obtaining estimates of the state of a nonlinear system is frequently solved by applying linear perturbation theory and using linear estimation theory to determine the perturbations. While this procedure is frequently satisfactory, many examples have been encountered in which second order effects are small but nonnegligible. The application of linear theory to nonlinear problems is essentially a means of approximating the mean and covariance of the a posteriori density. In this paper attention has been directed toward the development of approximations of these moments that include the influence of second order terms. The approximation involves the explicit assumption that the density is gaussian. This is not true for nonlinear system but it is felt that the first two moments will not be severly affected as long as the nonlinearity is small. Certainly, one would expect the accuracy of the moments containing the second order effects to be superior to the linear moments.

The immediate disadvantage of utilizing a quadratic perturbation theory lies in the increased number of system matrices that have to be determined. For linear systems the linear transition matrix is (nxn) and the observation matrix is(mxn). When the second order effects are included, there are n additional (nxn) plant matrices and m more (nxn) observation matrices to be calculated. The increase in computational requirements can be catastrophic for large n. This problem can be circumvented to an extent by establishing the equations in which the second order effects are significant and neglecting the insignificant effects in the other equations.

Although these results may require significantly more calculation, they do provide a systematic way for modifying linear estimation theory to include nonlinear effects. Numerical results for simple examples indicate that significant improvements in the response is possible when the nonlinearities are comparable to the noise. Thus, a sizable increase in the range of applicability of perturbative techniques may be provided.

REFERENCES

1. Sorenson, H.W., "Kalman Filtering Techniques", Chapter 5, Advances In Control Systems, Vol 3, edited by C.T. Leondos, Academic Press, New York, 1966.

2. Denham, W.F., S. Pines, "Sequential Estimation When Measurement Function Nonlinearity Is Comparable to Measurement Error", AIAA Journal, Vol.4, No 6, 1966, pp 1071-1076.

3. Bass, R.W., V.D. Norum, L. Schwartz, "Optimal Multichannel Nonlinear Filtering", Hughes Technical Report Number 55D 50064R, August 1965

4. Cox, H., "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems", IEEE Transactions on Automatic Control, AC-9, No 1, 1964, pp 5-12.

5. Sorenson, H.W., "A Nonlinear Perturbation Theory for Estimation and Control of Time-Discrete Stochastic Systems", PhD Dissertation. University of California, Los Angeles, 1966.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of California School of Engineering and Applied Science Los Angeles, California 90024	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

RECURSIVE FILTERING FOR SYSTEMS WITH SMALL BUT NONNEGLIGIBLE NONLINEARITIES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)
H.W. Sorenson
A. R. Stubberud

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July 1970	20	5

8a. CONTRACT OR GRANT NO. AF-AFOSR-699-67 b. PROJECT NO. 9749-01 c. 61102F 681304 d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR 70-1966 TR

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	Air Force Office of Scientific Research (SRM) 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

Linear estimation theory has been applied extensively to nonlinear systems by assuming that perturbations from a reference solution can be described by linear equations. As long as the second order (and higher) terms in the perturbation equations are negligible, linear estimation techniques have been found to yield satisfactory response. Many examples have been encountered in which the linear theory is not satisfactory, however, and it is to this situation that attention is directed here. Time-discrete systems in which the second order effects are small but nonnegligible are considered. Recursion relations for the conditional mean and covariance are developed. While these relations yield approximations to the true values of these moments, they are superior to the approximations provided by applying linear theory to a nonlinear system. Some results for a simple system are presented in which the response from linear and nonlinear filters is compared.

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification